12 Ramsey theory

Ramsey theory is the study of structure and of disorder. The main message of Ramsey theory is that *complete disorder is impossible*—any sufficiently large system, no matter how disordered, must contain within it some highly structured component. This general, highly unintuitive, philosophy manifests itself in topics as diverse as computer science, number theory, geometry, functional analysis, and, of course, graph theory, which is the topic we will mostly be focused on.

However, as Ramsey theory has connections to so many other areas of mathematics and beyond, we will also frequently pause to see how the results we have proved connect to these other fields. This is, in fact, how we begin the course, with perhaps the first-ever Ramsey-theoretic result, published by Issai Schur while Frank Ramsey was only fourteen years old.

12.1 Ramsey theory before Ramsey

Like many other people, Schur was interested in Fermat's last theorem, the statement that the equation $x^q + y^q = z^q$ has no non-trivial integer solutions x, y, z for any fixed $q \geqslant 3$, where a solution is *trivial* if $0 \in \{x, y, z\}$ and *non-trivial* otherwise.

Proving Fermat's last theorem is (very) hard, so let's start with something simpler. There are, of course, non-trivial integer solutions to the Pythagoras equation $x^2 + y^2 = z^2$. What if we change the equation slightly, to, say, $x^2 + y^2 = 3z^2$? After playing around with it for a bit, you might be tempted to conjecture that now, there are no non-trivial integer solutions.

This conjecture is indeed true, and there is a standard technique in number theory for proving such results. Namely, if there *were* some non-trivial solution $x, y, z \in \mathbb{Z}$ to the equation $x^2 + y^2 = 3z^2$, then there would also be a non-trivial[2] solution to the same equation modulo 3, namely the equation $x^2+y^2 \equiv 0 \pmod 3$. However, we know that that $1^2 \equiv 2^2 \equiv 1 \pmod 3$, and we can conclude that there *do not* exist non-trivial solutions modulo 3.

A similar argument can be used to prove that many other polynomial equations have no non-trivial integer solutions, and a general phenomenon called the *Hasse principle* very roughly says that in many instances, such a technique is guaranteed to work. So it is natural to wonder whether Fermat's last theorem can also be proved in this way. This is the question that motivated Schur[3], who proved that this technique *cannot* work for Fermat's last theorem.

Theorem 12.1 (Schur). *For any integer $q \geqslant 3$, there exists an integer $N = N(q)$ such that the following holds for any prime $p > N$. There exist non-zero $x, y, z \in \mathbb{Z}/p$ with*

$$x^q + y^q \equiv z^q \pmod p.$$

[2]One has to be a bit careful here, as a non-trivial solution over $\mathbb{Z}$ may become trivial in $\mathbb{Z}/3$. However, it is not hard to get around this issue, as one can argue that a *minimal* non-trivial solution over $\mathbb{Z}$ cannot have all three of x, y, z divisible by 3.

[3]In fact, the same question had motivated Dickson a few years earlier, and he was the first to prove Theorem 12.1. However, his technique used very messy casework and does not at all connect to Ramsey theory, so we won't discuss it any further.

As Schur himself realized, despite proving an important and impressive result in number theory, his proof used almost no number theory! He wrote “daß [Theorem 12.1] sich fast unmittelbar aus einem sehr einfachen Hilfssatz ergibt, der mehr der Kombinatorik als der Zahlentheorie angehört.”[4] This Hilfssatz is the following.

Theorem 12.2 (Schur)**.** *For any positive integer q, there exists an integer $N = N(q)$ such that the following holds. If $[\![N]\!]$ is colored in q colors, then there exist $x, y, z \in [\![N]\!]$, all receiving the same color, such that $x + y = z$.*

Recall that we use the notation $[\![N]\!] := \{1, \ldots, N\}$. We also now start to use the terminology of *coloring*. By a coloring of $[\![N]\!]$ with q colors, we just mean a partition of $[\![N]\!]$ into q sets $A_1, \ldots, A_q$, where we think of the elements of A_1 as receiving a first color, the elements of A_2 as receiving some second, distinct, color, and so on. We will also frequently use the shorthand *monochromatic* for “receiving the same color”, so the conclusion of Theorem 12.2 could also be stated as the existence of a monochromatic solution to $x + y = z$.

As Schur wrote, the derivation of Theorem 12.1 from Theorem 12.2 is almost immediate, but as it requires a few ideas from number theory and group theory, we will defer it for the moment. Let us first see how to prove Theorem 12.2. Schur proved Theorem 12.2 directly, but the modern, Ramsey-theoretic, perspective is to reduce Theorem 12.2 to an even more combinatorial lemma, which we now state.

Lemma 12.3. *For any positive integer q, there exists an integer $N = N(q)$ such that the following holds. If the edges of the complete graph K_N are q-colored, then there exists a monochromatic triangle.*

Proof. We will actually prove something stronger, namely an explicit upper bound on $N(q)$; we will show that $N(q) = 3q!$ satisfies the desired condition. We proceed by induction on q.

The base case $q = 1$ is immediate. We are claiming that any 1-coloring of the edges of K_N, where $N = 3 \cdot 1! = 3$, contains a monochromatic triangle. But as there is only one color, and the complete graph we are “coloring” is itself a triangle, this is certainly true.

For the inductive step, suppose the result is true for $q-1$, i.e. that any $(q-1)$-coloring of $E(K_{3(q-1)!})$ contains a monochromatic triangle. Fix a q-coloring of $E(K_N)$, where $N = 3q!$, and let v be any vertex of K_N. v is incident to $N - 1$ edges, each of which receives one of q colors. Therefore, by the pigeonhole principle, there is some color, say red, which appears on at least

$$\left\lceil \frac{N-1}{q} \right\rceil = \left\lceil \frac{3q!-1}{q} \right\rceil = \left\lceil 3(q-1)! - \frac{1}{q} \right\rceil = 3(q-1)!$$

edges incident to v. Let R denote the set of endpoints of these red edges, and consider the coloring restricted to R. If there is any red edge appearing in R, then it forms a red triangle together with v, and we are done. If not, then R is a set of at least $3(q-1)!$ vertices that are colored by at most $q-1$ colors, and we can find a monochromatic triangle in R by the inductive hypothesis. In either case we are done. □

[4] “that [Theorem 12.1] follows almost immediately from a very simple lemma, which belongs more to combinatorics than to number theory.”

With Lemma 12.3 in hand, the proof of Theorem 12.2 is almost immediate. All we need to do is to translate the number-theoretic coloring into a graph-theoretic coloring.

Proof of Theorem 12.2. Let $N(q) = 3q!$ be chosen so that Lemma 12.3 holds. We are given a q-coloring χ of $[\![N]\!]$, which we convert to a q-coloring $\hat{\chi}$ of $E(K_N)$ as follows. Identify the vertices of K_N with $[\![N]\!]$, and then color an edge ab, where $1 \leqslant a < b \leqslant N$, according to the color of $b - a \in [\![N]\!]$ in χ.

As $\hat{\chi}$ is a q-coloring of $E(K_N)$, by Lemma 12.3, there is a monochromatic triangle in $\hat{\chi}$. Let the vertices of this triangle be a, b, c, where $a < b < c$. Let $x = b - a, y = c - b$, and $z = c - a$, and note that these satisfy $x + y = z$. Finally, note that they all receive the same color under χ, since $\chi(x) = \hat{\chi}(ab), \chi(y) = \hat{\chi}(bc)$, and $\chi(z) = \hat{\chi}(ac)$, and we assumed that a, b, c is a monochromatic triangle under $\hat{\chi}$. □

This completes the combinatorial part of Schur's work. For completeness, let's see how to derive Theorem 12.1 from Theorem 12.2. As this topic is somewhat outside the main narrative of the class, it will not be covered in lecture.

Deduction of Theorem 12.1 from Theorem 12.2

Proof of Theorem 12.1. Let $N = N(q)$ be as in Theorem 12.2, and fix a prime $p > N$. We recall the well-known fact that the set $\Gamma := \{x^q : 0 \neq x \in \mathbb{Z}/p\}$ forms a subgroup of the multiplicative group $(\mathbb{Z}/p)^\times$, and the index of this subgroup is at most[†] q. Therefore, there are at most q cosets of Γ which partition the non-zero elements of $\mathbb{Z}/p$. By identifying the non-zero elements of $\mathbb{Z}/p$ with $[\![p-1]\!] \supseteq [\![N]\!]$, we obtain a q-coloring of $[\![N]\!]$ according to these cosets.

Now, by Theorem 12.2, there must exist monochromatic $a, b, c \in [\![N]\!]$ such that $a + b = c$. As these three numbers receive the same color, they must lie in some single coset $\alpha\Gamma$ of Γ, for some $\alpha \in (\mathbb{Z}/p)^\times$. By the definition of Γ, this means that we can write

$$a \equiv \alpha x^q \pmod p, \qquad b \equiv \alpha y^q \pmod p, \qquad c \equiv \alpha z^q \pmod p,$$

for some non-zero $x, y, z \in \mathbb{Z}/p$. The equation $a + b = c$ remains true when we reduce it mod p, so we conclude that

$$\alpha x^q + \alpha y^q \equiv \alpha z^q \pmod p.$$

As α is invertible in $\mathbb{Z}/p$, and as $x, y, z \neq 0$, we obtained the desired non-trivial solution $x^q + y^q \equiv z^q \pmod p$. □

[†] More precisely, the index is exactly $\gcd(q, p-1)$.

13 Classical Ramsey numbers

13.1 Ramsey's theorem and upper bounds on Ramsey numbers

While Schur's theorem can be seen as an early example of Ramsey theory, the theory did not really get going until Frank Ramsey's pioneering work in 1929. Ramsey's theorem, as it is now called, is a generalization of Lemma 12.3 from triangles to arbitrary cliques.

Theorem 13.1 (Ramsey)**.** *For all positive integers k, q, there exists an integer $N = N(k, q)$ such that the following holds. If the edges of the complete graph K_N are q-colored, then there exists a monochromatic K_k, that is, k vertices such that all the $\binom{k}{2}$ edges between them receive the same color.*

Given this theorem, which we will shortly prove, we can make a definition that will be central for much of the rest of the course.

Definition 13.2. Given positive integers k, q, the *q-color Ramsey number of K_k*, denoted $r(k; q)$, is the least N such that the conclusion of Theorem 13.1 is true. That is, $r(k; q)$ is the minimum integer N such that every q-coloring of $E(K_N)$ contains a monochromatic K_k.

In case $q = 2$, we usually abbreviate $r(k; 2)$ as simply $r(k)$, and usually refer to the 2-color Ramsey number as simply the *Ramsey number*.

In this language, Theorem 13.1 can equivalently be stated as saying that $r(k; q) < \infty$ for all k, q. In fact, for much of this course, we will be interested not just in the fact that such Ramsey numbers are finite, but in quantitative estimates on how large they are.

For now, let's focus on the case $q = 2$. Ramsey's original proof of Theorem 13.1 showed that $r(k) \leqslant k!$ for all k. But a few years later, a different proof was found by Erdős and Szekeres, in another foundational paper of the field. In order to present their proof, we need to define a slightly more general notion of Ramsey number.

Definition 13.3. Given positive integers k, ℓ, we denote by $r(k, \ell)$ the *off-diagonal Ramsey number*, defined to be the least N such that every 2-coloring of $E(K_N)$ with colors red and blue contains a red K_k or a blue K_ℓ.

Note that $r(k, \ell) = r(\ell, k)$ as the colors play symmetric roles, and that $r(k) = r(k, k)$.

Theorem 13.4 (Erdős–Szekeres)**.** *For all positive integers k, ℓ, we have*

$$r(k, \ell) \leqslant \binom{k + \ell - 2}{k - 1}.$$

In particular, we have

$$r(k) \leqslant \binom{2k - 2}{k - 1} < 4^k.$$

Proof. We proceed by induction on $k + \ell$, with the base case[5] $k = 1$ or $\ell = 1$ being trivial. For the inductive step, the key claim is that the following inequality holds:

$$r(k, \ell) \leqslant r(k - 1, \ell) + r(k, \ell - 1). \tag{4}$$

[5]If you don't like starting the induction with $k = 1$—what does a monochromatic K_1 mean, exactly?—you should convince yourself that the base case $k = 2$ or $\ell = 2$ also works.

To prove (4), fix a red/blue coloring of $E(K_N)$, where $N = r(k-1,\ell)+r(k,\ell-1)$, and fix some vertex $v \in V(K_N)$. Suppose for the moment that v is incident to at least $r(k-1,\ell)$ red edges, and let R denote the set of endpoints of these red edges. By definition, as $|R| \geqslant r(k-1,\ell)$, we know that R contains a red K_{k-1} or a blue K_ℓ. In the latter case we have found a blue K_ℓ (so we are done), and in the former case we can add v to this red K_{k-1} to obtain a red K_k (and we are again done).

So we may assume that v is incident to fewer than $r(k-1,\ell)$ red edges. By the exact same argument, just interchanging the roles of the colors, we may assume that v is incident to fewer than $r(k,\ell-1)$ blue edges. But then the total number of edges incident to v is at most

$$(r(k-1,\ell)-1)+(r(k,\ell-1)-1) = N-2,$$

which is impossible, as v is adjacent to all $N-1$ other vertices. This is a contradiction, proving (4).

We can now complete the induction. By (4) and the inductive hypothesis, we find that

$$r(k,\ell) \leqslant r(k-1,\ell)+r(k,\ell-1) \leqslant \binom{(k-1)+\ell-2}{(k-1)-1} + \binom{k+(\ell-1)-2}{k-1} = \binom{k+\ell-2}{k-1},$$

where the final equality is Pascal's identity for binomial coefficients. □

A similar argument works when the number of colors is more than 2. If we denote by $r(k_1,\ldots,k_q)$ the *off-diagonal multicolor Ramsey number* (defined in the natural way), we obtain the following generalization of Theorem 13.4, which you will prove on the homework.

Theorem 13.5. *For all positive integers q and $k_1,\ldots,k_q$, we have*

$$r(k_1,\ldots,k_q) \leqslant \binom{k_1+\cdots+k_q-q}{k_1-1,\ldots,k_q-1},$$

where the right-hand side denotes the multinomial coefficient. In particular,

$$r(k;q) < q^{qk}.$$

13.2 Lower bounds on Ramsey numbers

The Erdős–Szekeres bound, Theorem 13.4, gives us the upper bound $r(k) < 4^k$, which improves on Ramsey's earlier bound of $r(k) \leqslant k!$. To understand how good this bound is, we would like to obtain some *lower bounds* on $r(k)$.

Thinking about the definition of Ramsey numbers, we see that proving a lower bound of $r(k) > N$ boils down to exhibiting a 2-coloring of $E(K_N)$ with no monochromatic K_k. Perhaps the simplest such coloring is the *Turán coloring*, which proves the following result.

Proposition 13.6. *For any positive integer k, we have $r(k) > (k-1)^2$.*

Proof. Let $N = (k-1)^2$. We split the vertex set of K_N into $k-1$ parts, each of size $k-1$. We color all edges within a part red, and all edges between parts blue. The red graph is a disjoint union of $k-1$ copies of K_{k-1}, so there is certainly no red K_k. On the other hand, as there are only $k-1$ parts, the pigeonhole principle implies that any set of k vertices must include two vertices in one part; these two vertices span a red edge, and thus there is no blue K_k either. □